Some econometric applications of the exact distribution of the ratio of two quadratic forms in normal variates

Palm, F.C.; Sneek, J.M.

1981

SOME ECONOMETRIC APPLICATIONS OF THE EXACT DISTRIBUTION OF THE RATIO OF TWO QUADRATIC FORMS IN NORMAL VARIATES

F.C. Palm
J.M. Sneek

Researchmemorandum 1981-18 Sept. 1981

Some Econometric Applications of the Exact Distribution of the Ratio of Two Quadratic Forms in Normal Variates.

F.C. Palm and J.M. Sneek*

Abstract

Ratios of quadratic forms in normal variates arise in many econometric and statistical applications. Their exact distribution can be computed using e.g. a procedure due to Imhof (1961).

In this paper two examples arising in dynamic models will be considered. First, the distribution of a test of linear restrictions on the coefficients of a regression model with autocorrelated errors will be analyzed. Second, the distribution of the sample autocorrelations of a series generated by an autoregressive - integrated - moving average model will be investigated.
In both cases, the exact distribution will be compared with the (approximate) distributions used in applied work.
Some remarks on the usefulness of the exact distribution of quotients of quadratic forms in normal variates conclude the paper.

Keywords: ratio of quadratic form, regression model, F-statistic
autocorrelation coefficient, ARIMA model

* Address: Economische Faculteit, Vrije Universiteit,
Postbus 7161, 1007 MC Amsterdam, The Netherlands.

Some Econometric Applications of the Exact Distribution of the Ratio of Two Quadratic Forms in Normal Variates.

F.C. Palm and J.M. Sneek

August 1981

Comments welcome

1. Introduction

Ratios of quadratic forms in normal variates arise in many econometric and statistical applications. Perhaps the most common examples are the F-statistic used to test linear restrictions on the coefficients in the linear regression and analysis of variance models, the Durbin-Watson test and the Von Neumann ratio for testing against first order autocorrelation in the disturbances of the linear regression model and similar expressions for testing against higher order disturbance autocorrelation. We should also mention the multiple correlation coefficient, R^2, in regression models.

In the Box-Jenkins (1970) approach to time series analysis, the estimated autocorrelation coefficients, which are quotients of quadratic forms in the observations, are used to identify the model for the time series. There are several ways to handle the problem involved with the distribution of ratios of quadratic forms in normal variates (RQFNV). First, in some cases, the exact distribution is known. For instance, in the standard linear regression model, the numerator and denominator of the F-statistic for linear restrictions are chi-square distributed and independent and therefore the quotient is F-distributed. Second, quite often the large sample distribution of the RQFNV has been established. Box and Jenkins (1970) use the asymptotic distribution of the autocorrelation function in order to identify an autoregressive - integrated - moving average (ARIMA) model for the series. Sometimes, when the usual 'F-statistic' for testing linear restrictions on the regression coefficients does not have an F-distribution, a slight transformation of the statistic can be shown to have a large sample chi-square distribution. Third, one can use other approximations to the exact distribution. For instance, Durbin and Watson (1950) computed lower and upper bounds for the d-statistic which make their test independent of the values of the explanatory variables.

Kiviet (1979) gives lower and upper bounds for the usual t- and F-statistics when the disturbances in the regression model are autocorrelated. In an attempt to improve the model identification by means of the sample autocorrelations, Anderson (1979) gives formulae for the expectations of the sample variance and covariances for series generated by ARIMA models. Similarly, De Gooijer (1980) investigates the exact moments of the sample autocorrelations for some time series models.

The fourth possibility, which we deal with in the paper, consists in computing the exact distribution of the RQFNV. Imhof (1961) presents a procedure for computing the distribution of RQFNV's. His procedure has been used by several authors. We shall mention a few among them. Koerts and Abrahamse (1969) compute the exact distribution of the BLUS test for first order disturbance autocorrelation and of the multiple correlation coefficient, R^2, in a regression model.

Savin and White (1977) compute the exact distribution of the bounding random variables d_L and d_U of the Durbin-Watson test for a sample size up to 200 and up to 20 regressors. Farebrother (1980) gives the exact distribution for the minimal bound when there is no intercept in the regression with T up to 200 and up to 21 regressors.

In this paper, we use Imhof's procedure to compute the exact distribution of the RQFNV arising when the usual F-statistic is used in a regression model with autocorrelated disturbances. We also give results for the distribution of the 'F-statistic' when a generalized least squares estimator is used with a disturbance covariance matrix that differs from the true disturbance covariance matrix. We compare the exact distribution with the lower and upper bounds derived by Kiviet (1979). Results for the exact confidence region of the regression coefficients, when OLS estimation is applied to a model with first order autocorrelated errors, have also been obtained by Kiviet (1977). Imhof's procedure is also used to obtain the distribution of sample autocorrelations for ARIMA-models. The results are compared with some approximations to the exact distribution.

Our results show that in both cases, it may be worthwhile to compute exact probabilities rather than to rely on computationally less expensive approximations. Furthermore, the costs of computing the exact distribution seem to be reasonable. In section 2, we briefly discuss Imhof's procedure and indicate how it can be applied to compute tail area probabilities for RQFNV's. Section 3 is devoted to the distribution of tests for linear restrictions on the regression coefficients, when disturbance autocorrelation is present. In section 4, we present some results on the distribution of the sample autocorrelations for ARIMA-models. The reader who is interested in the results on the sample autocorrelations only may skip section 3. Section 5 provides a brief summary and contains some concluding remarks.

2. Computing the distribution of a RQFNV

In this section, we briefly review the method of computing the tail area probability

$$\Pr\left[\frac{\underline{\eta}'\, A\, \underline{\eta}}{\underline{\eta}'\, B\, \underline{\eta}} \leq q\right], \qquad (2.1)$$

where A and B are symmetric matrices, and B is semi-definite, $\underline{\eta}$ is a normally distributed stochastic vector of length n with mean $E\underline{\eta}$ and covariance matrix $\Sigma = CC'$ and q is some real number. If Σ is non-singular, expression (2.1) can be written as:

$$\Pr\,[\underline{\eta}'\,(A - qB)\,\underline{\eta} \leq 0] \;=\; \Pr\,(\sum_{j=1}^{n} \lambda_j\, \underline{v}_j^2 \leq 0)\,, \qquad (2.2)$$

where the λ_j's are the characteristic roots of $C'[A - qB]C$ and the $\underline{v}_j$'s are independent normally distributed variables with mean μ_j depending on $E\underline{\eta}$ and Σ and with unit variance. We can write

$\sum_{j=1}^{n} \lambda_j\, \underline{v}_j^2 = \sum_{r=1}^{m} \tilde{\lambda}_r\, \underline{\chi}^2\,(h_r\,,\,\nu_r^2)$, where the $\tilde{\lambda}_r$'s are non-zero distinct roots, $m \leq n$, h_r is the number of λ_j's equal to $\tilde{\lambda}_r$ and $\underline{\chi}^2\,(h_r\,,\,\nu_r^2)$ has a non-central chi-square distribution with h_r degrees of freedom and non-centrality parameter $\nu_r^2 = \Sigma\mu_j^2$, where summation runs over those j's for which $\lambda_j = \tilde{\lambda}_r$. Imhof (1961) shows that the probability in (2.2) can be obtained from

$$\Pr\,[\sum_{r=1}^{m} \tilde{\lambda}_r\, \underline{\chi}^2\,(h_r\,,\,\nu_r^2) \leq x] = \tfrac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \frac{\sin\,\Theta(u)}{u\,\rho\,(u)}\,du, \quad (2.3)$$

where

$$\Theta(u) = \tfrac{1}{2}\sum_{r=1}^{m}\,[h_r\,\mathrm{arctg}\,(\tilde{\lambda}_r\,u) + \frac{\nu_r^2\,\tilde{\lambda}_r\,u_r}{1 + \tilde{\lambda}_r^2\,u^2}] - \tfrac{1}{2}x\,u$$

$$\rho(u) = \prod_{r=1}^{m}\,(1 + \tilde{\lambda}_r^2\,u^2)^{h_r/4}\;\exp\;\tfrac{1}{2}\left[\sum_{r=1}^{m} \frac{(\nu_r\,\tilde{\lambda}_r\,u)^2}{1+\tilde{\lambda}_r^2\,u^2}\right].$$

Expression (2.3) can be derived by considering the following inversion formula for characteristic functions:

$$F_{\underline{w}}\,(x) = \tfrac{1}{2} - \frac{1}{\pi}\int_0^{\infty} \frac{1}{t}\,\mathrm{Im}\;\{e^{-itx}\,f(t)\}\,dt. \qquad (2.4)$$

In (2.4), $F_{\underline{w}}\,(x)$ denotes the distribution function of the random variable $\underline{w}$, $f(t)$ is the characteristic function of $\underline{w}$ and Im $\{z\}$ is the imaginary part of z. Although it is possible to get analytical expressions for (2.3) in some special cases, it is usually more convenient to integrate (2.3) numerically.

When the covariance matrix Σ is singular, say of rank $p < n$, one can write $\underline{\eta} = \gamma + L\underline{v}$, with $\gamma = E\underline{\eta}$ and L being an n x p matrix such that $\Sigma = LL'$, $\underline{v} \sim N(0\,,\,I)$ is a p x 1 vector of random variables. In this case, the probability in (2.1) is equal to

$$\Pr\left[(\underline{v}'(L'AL - qL'BL)\underline{v} + 2\gamma'(AL - qBL)\underline{v} \leq c)\right] =$$

$$\Pr\left[\sum_{j=1}^{p}(\lambda_j \underline{w}_j^2 + 2a_j \underline{w}_j) \leq c\right] = \Pr\left[\sum_{j=1}^{s}\lambda_j\left(\underline{w}_j + \frac{a_j}{\lambda_j}\right)^2 + 2\sum_{j=s+1}^{p} a_j \underline{w}_j \leq \bar{c}\right] =$$

$$\Pr\left[\sum_{r=1}^{s'}\tilde{\lambda}_r \underline{\chi}^2(h_r, \nu_r^2) + 2\sum_{j=s+1}^{p} a_j \underline{w}_j \leq \bar{c}\right] =$$

$$\Pr\left[\underline{u} + \underline{y} \leq \bar{c}\right], \qquad (2.5)$$

with $c = -\gamma'(A - qB)\gamma$,

$\bar{c} = c + \sum_{j=1}^{s}\frac{a_j^2}{\lambda_j}$,

$\underline{w}_j$ is defined through an orthogonal transformation $\underline{v} = P\underline{w}$, such that the symmetric matrix $L'(A - qB)L$ diagonalizes ,

a_j is the j-th element of $\gamma'(AL - qBL)P$.

Furthermore, it is assumed that the first s values of λ_j are the only non-zero values in the set $\{\lambda_1, \ldots, \lambda_n\}$, s' denotes the number of distinct non-zero characteristic values λ_j.
The non-centrality parameter ν_r^2 equals $\Sigma\left[\frac{a_j}{\lambda_j}\right]^2 = \frac{1}{\tilde{\lambda}_r^2}\Sigma a_j^2$, where summation runs over those values of j, for which $\lambda_j = \tilde{\lambda}_r$.
As the characteristic function of $\underline{y}$, $\varphi_{\underline{y}}(t) = \exp\left[-2t^2\sum_{j=s+1}^{p} a_j^2\right]$, is a real function, the probability in (2.5) can be obtained by applying formulae (2.4) and (2.3) to yield

$$\Pr(\underline{u} + \underline{y} \leq \bar{c}) = \tfrac{1}{2} - \frac{1}{\pi}\int_0^\infty \exp\left[-\frac{u^2}{2}\sum_{j=s+1}^{p} a_j^2\right]\frac{\sin\Theta(u)}{u\,\rho(u)}\,du, \qquad (2.6)$$

where $\Theta(u) = \tfrac{1}{2}\sum_{r=1}^{s}\left[h_r \operatorname{arctg}(\tilde{\lambda}_r u) + \frac{\nu_r^2\tilde{\lambda}_r u}{1 + \tilde{\lambda}_r^2 u^2}\right] - \tfrac{1}{2}\bar{c}u$,

$$\rho(u) = \prod_{r=1}^{s'}(1 + \tilde{\lambda}_r^2 u^2)^{h_r/4}\exp\left\{\tfrac{1}{2}\sum_{r=1}^{s'}\frac{(\nu_r\tilde{\lambda}_r u)^2}{1 + \tilde{\lambda}_r^2 u^2}\right\} .$$

Obviously, (2.3) is a special case of (2.6). Expressions (2.3) and (2.6) will be used in sections 3 and 4 in order to evaluate the distribution of some RQFNV's.
A Fortran-version of the computer programs given by Koerts and Abrahamse (1969)

will be used. Some minor modifications similar to those of Farebrother (1980) have been made to assure the convergence of the numerical integration. The truncation and integration errors have been fixed at resp. .0001 in section 3 and .001 in section 4.

3. Testing linear restrictions on the regression coefficients of the linear regression model

3.1 The model and the test statistics

Consider the following linear regression model:

$$\underset{T\times 1}{\underline{y}} = \underset{T\times k}{X}\,\underset{k\times 1}{\beta} + \underset{T\times 1}{\underline{u}} \quad , \tag{3.1}$$

where $\underline{y}$ is a vector of random variables,

X is a matrix of regressors, with full rank,

β is a vector of regression coefficients,

$\underline{u}$ is a vector of disturbances, assumed to be normally distributed with mean zero and covariance matrix Ω, symmetric and positive definite, $\Omega = CC'$ with C being lower triangular. The disturbances $\underline{u}$ and X are independent.

With known matrix Ω, application of ordinary least squares (OLS) to the system (3.1) after premultiplication of (3.1) by C^{-1} yields the generalized least squares (GLS) estimator of β which is best, linear and unbiased and is identical with the OLS estimator if and only if $\Omega = X\Gamma X' + Z\Theta Z' + \sigma^2 I$, where Γ, Θ and σ^2 are arbitrary and Z is a matrix such that $X'Z = 0$ (see Rao (1967)). The F-statistic for testing a set of m linear restrictions on β, $R\beta = r$, with R being an $m \times k$ matrix of rank $m \leq k$ and r an $m \times 1$ vector, is given by

$$\underline{F} = \frac{(K\underline{y} - r)'\, Q\, (K\underline{y} - r)}{\underline{y}' M\, \underline{y}} \quad , \tag{3.2}$$

where
$$K = R\,(X'\,\Omega^{-1}X)^{-1}\,X'\,\Omega^{-1} \quad ,$$

$$Q = [R\,(X'\,\Omega^{-1}\,X)^{-1}\,R']^{-1} / m \quad ,$$

$$M = [\Omega^{-1} - \Omega^{-1}\,X\,(X'\,\Omega^{-1}\,X)^{-1}\,X'\,\Omega^{-1}] / T - k \quad .$$

Under the null hypothesis H_0: $R\beta = r$, the test-statistic in (3.2) has an $F(m, T\text{-}k)$ - distribution. Furthermore, under the alternative hypothesis, H_1: $R\beta \neq r$, expression (3.2) is distributed as a non-central F- distribution with m and T-k degrees of freedom and non-centrality parameter $\nu = [R\beta - r]'\, Q\, [R\beta - r]$.

Obviously, when $\Omega = \sigma^2 I$, these results hold true for expression (3.2) appropriately specialized to

$$\tilde{\underline{F}} = \frac{[R\,(X'X)^{-1}\,X'\underline{y} - r]'[R\,(X'X)^{-1}R']^{-1}[R\,(X'X)^{-1}X'\underline{y}-r]\;/\;m}{\underline{y}'\,[I - X(X'X)^{-1}\,X']\underline{y}\;/\;T-k} \,. \qquad (3.3)$$

When $\Omega \neq \sigma^2 I$, the statistic $\tilde{\underline{F}}$ in (3.3) becomes

$$\tilde{\underline{F}} = \frac{[p + K_1\,\underline{\eta}]'\,Q_1\,[p + K_1\,\underline{\eta}]}{\underline{\eta}'\,M_1\,\underline{\eta}} \quad , \qquad (3.4)$$

with
$$\begin{aligned} \underline{\eta} &= C^{-1}\,\underline{u} \quad \text{being distributed as } N\,(0\,,I) \\ p &= R\beta - r \\ K_1 &= R\,(X'X)^{-1}\,X'\,C \\ Q_1 &= [R\,(X'X)^{-1}\,R']^{-1}\,/\,m \\ M_1 &= C'\,[I - X\,(X'X)^{-1}\,X']\,C\,/\,T-k\,. \end{aligned}$$

Expression (3.4) does not have an F-distribution, as it is no longer a quotient of two independent chi-square distributed random variables. Under H_0 : $p = 0$, expression (3.4) is a ratio of quadratic forms in the normal variates $\underline{\eta}$, with $A = K_1'\,Q_1\,K_1$ and $B = M_1$, where neither A nor B are idempotent.
For $p = 0$, the computation of the exact distribution of $\tilde{\underline{F}}$ in (3.4) requires integration of expression (2.3) appropriately specialized to take into account the nullity of the noncentrality parameters ν_r under H_0 . The quantities $\tilde{\lambda}_r$ are the different non-zero characteristic roots of $A - qB = K_1'\,Q_1\,K_1 - q\,M_1$, with q being the argument in the distribution function, and h_r is the multiplicity of the non-zero characteristic roots $\tilde{\lambda}_r$. Notice that Imhof's procedure could also be applied to get the distribution of $\underline{F}$ in (3.2) under $H_0 : R\beta = r$. In that case, there would be only two different values of $\tilde{\lambda}_r$, equal to respectively $1/m$ and $-q/T-k$ with the corresponding values of h_r being equal to respectively m and $(T-k)$.

Under H_1 : $R\beta \neq r$ or equivalently $p \neq 0$, the distribution of $\tilde{\underline{F}}$ in (3.4) can be computed along the lines of the second part of section 2. The probability that $\tilde{\underline{F}}$ in (3.4) is smaller than q can be written as:

$$\Pr(\tilde{\underline{F}} \leq q) = \Pr(\underline{\eta}'\,[K_1'\,Q_1\,K_1 - q\,M_1]\underline{\eta} + 2\,\underline{\eta}'\,K_1'\,Q_1\,p + p'\,Q_1\,p \leq 0)\,. \qquad (3.5)$$

There exists an orthogonal matrix P that diagonalizes the symmetric matrix $[K_1'\,Q_1\,K_1 - q\,M_1]$. Define $\underline{w} = P^{-1}\,\underline{\eta}$. Then $\underline{w}$ is distributed as $N\,(0\,,I)$ and expression (3.5) becomes:

$$\Pr(\underset{\sim}{\tilde{F}} \leq q) = \Pr\left(\sum_{j=1}^{T} \lambda_j \underset{\sim}{w}_j^2 + 2 \sum_{j=1}^{T} \mu_j \underset{\sim}{w}_j + p' Q_1 p \leq 0\right) \quad , \tag{3.6}$$

where the λ_j's are the characteristic roots of $[K_1' Q_1 K_1 - q M_1]$ and μ_j is the j-th element of the $1 \times T$ vector $\mu' = p' Q_1' K_1' P$. Assuming the first s values of λ_j to be the only non-zero eigen-values , the probability in (3.6) can be computed in the same way as the expression (2.6) is computed.

Obviously under H_0 the distribution of $\underset{\sim}{\tilde{F}}$ is a special case of (3.6). Notice also that, under both H_0 and H_1 , the exact distribution of the statistic $\underset{\sim}{F}$ in (3.2) can be computed along the lines discussed in section 2.

Finally, we shall compute the exact distribution of the so-called F-statistic obtained when the GLS-estimator is applied to the model (3.1) using an approximation $\hat{\Omega} \neq \Omega$. In this case, the 'F-statistic' becomes

$$\underset{\sim}{\hat{F}} = \frac{[p + K_2 \underset{\sim}{\eta}]' \, Q_2 \, [p + K_2 \underset{\sim}{\eta}]}{\underset{\sim}{\eta}' M_2 \underset{\sim}{\eta}} \quad , \tag{3.7}$$

where
$$K_2 = R\,(X' \hat{\Omega}^{-1} X)^{-1} X' \hat{\Omega}^{-1} C$$
$$Q_2 = [R\,(X' \hat{\Omega}^{-1} X)^{-1} R'] / m$$
$$M_2 = C' [\hat{\Omega}^{-1} - \hat{\Omega}^{-1} X (X' \hat{\Omega}^{-1} X) \hat{\Omega}^{-1}] C / T - k \; .$$

Before we compute the distribution of $\underset{\sim}{\hat{F}}$, we have to distinguish between the situation, where $\hat{\Omega}$ is a given non-stochastic matrix and that where $\underset{\sim}{\hat{\Omega}}$ has been estimated and is stochastic. With a stochastic matrix $\underset{\sim}{\hat{\Omega}}$, the distribution of $\underset{\sim}{\hat{F}}$ can be obtained in two stages, i.e. first get the distribution of $\underset{\sim}{\hat{F}}$ conditionally on a value $\hat{\Omega}$ of $\underset{\sim}{\hat{\Omega}}$ using e.g. Imhof's procedure, second integrate with respect to the marginal distribution of $\underset{\sim}{\hat{\Omega}}$.

It should be obvious that the two-stage procedure is feasible only when $\hat{\Omega}$ is a function of a very small number of parameter estimates, as the integration has to be done over the parameter space for $\hat{\Omega}$. In addition, the distribution of $\underset{\sim}{\hat{\Omega}}$ has to be known (perhaps approximately from simulations or from the large sample theory). In order to get a good approximation of the distribution of $\underset{\sim}{\hat{F}}$ with estimated $\underset{\sim}{\hat{\Omega}}$, using numerical integration , the grid for the values of the elements in $\hat{\Omega}$ should depend on the density of the estimator for these elements [e.g. use 'importance sampling', see Hammersley and Handscomb (1964)]. The characteristic roots of $A - qB$ associated with (3.7) have to be calculated for every value of the elements of $\hat{\Omega}$. In conclusion, when $\underset{\sim}{\hat{\Omega}}$ is a stochastic matrix, the evaluation of the exact distribution of $\underset{\sim}{\hat{F}}$ in (3.7) will be very cumbersome, except in special situations such as e.g. when the disturbances in (3.1) are generated by a first order autoregressive process. Usually, one will have to be satisfied with the large sample Wald test, $m\,\underline{F}$, being $\chi^2(m)$-distributed, where $\underline{F}$ is the quantity given in (3.2) with Ω replaced by a consistent estimate.

When $\hat{\Omega}$ is a nonstochastic matrix, the distribution of $\hat{\underline{F}}$ in (3.7) can be computed using Imhof's procedure. Note that under H_0, a scaling factor σ^2 in Ω has no influence on the distribution of $\hat{\underline{F}}$. When only a scaling factor σ^2 in Ω has to be estimated, it is preferable to parametrize the process for $\underline{u}$ in (3.1) as $N(0, \sigma^2\,\Omega)$, because this leads to considerable simplifications of the computations. It is obvious that under H_0, expression (3.7) does not depend on the estimate of σ^2. Under H_1, an overestimate of σ^2 results in underestimation of the power function and vice versa, as can be seen from expression (3.3) where the non-centrality parameter varies proportionally to σ^{-2}.

3.2. Some numerical results for the F-statistic

In this section we will give some results for the F-statistic in the general linear regression model (3.1). For the model with first order autoregressive disturbances, Kiviet (1977) has computed the exact tail area of the "F-" and "t-test" on regression coefficients estimated by OLS. Kiviet (1979) has obtained lower and upper bounds for the probability of an error of type I if one uses the usual t- and F-statistic in (3.4), when the disturbances are generated by simple ARMA-schemes. The bounds given by Kiviet (1979) are independent of the regressor matrix X and the restriction matrix R. From his results, it becomes clear that the conclusions for the usual t- and F-statistic are heavily affected, when the autocorrelation in the disturbances is incorrectly taken into account. But the range of values for which these tests are inconclusive is quite large. Therefore, the exact distribution of these statistics is expected to give additional insight into their properties and their sensitivity with respect to departures from the assumptions on the disturbance correlations.

In this section, we extend the analysis of the exact distribution to the situations, where the disturbance autocorrelation is taken into account by using GLS. This procedure is formally equivalent to OLS-estimation applied to the transformed model in which the disturbances then follow an ARMA-process.We also compare the exact distribution with the bounds reported in Kiviet (1979). As the exact distribution depends on the characteristics of the regressor matrix, it is sensible to use a general linear regression model with a 'typically economic' regressor matrix (see e.g. Dubbelman, Louter and Abrahamse [1978]). In this way, we hope to obtain results for models similar to those used in applied econometrics. We have chosen the data for the Netherlands given by Theil (1971, p. 102) in Table 3.1 , where the regressors consist of a constant term, the logarithm of the income per capita and the relative price of textiles for the period 1923-1939.

In Table 1, the probability $\Pr(\hat{\underline{F}} \leq q)$ is tabulated for the general linear regression model (3.1), where the disturbance term $\underline{u}$ is generated by a first order autoregressive process $\underline{u}_t = \rho\, \underline{u}_{t-1} + \underline{\varepsilon}_t$, with $\underline{\varepsilon} \sim N(0, I)$. Different values of ρ have been used. The sample size is always $T = 17$. The value for r reported in column 6 in Table 1 has been used in the matrix $\hat{\Omega}$, treated as non-stochastic ,

$$\hat{\Omega} = \begin{bmatrix} 1 & r & r^2 & \cdots & r^{T-1} \\ r & 1 & & & \cdot \\ r^2 & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \\ \cdot & & & & \\ r^{T-1} & & \cdots & & 1 \end{bmatrix} . \tag{3.8}$$

As we evaluate $\Pr(\hat{\underline{F}} \leq q)$ only under H_0 , the scaling factors in Ω , $\sigma^2 / 1 - \rho^2$, and in $\hat{\Omega}$, $\hat{\sigma}^2 / 1 - r^2$, can be ignored. The different sets of exclusion restrictions considered are given in column 7 of Table 1. In column 8, we report the number of seconds (for the two corresponding rows) that an execution takes on a CDC CYBER 170-750 computer. Truncation and integration errors have been fixed at .0001 .

Only a few examples of restrictions have been considered. The values for q in Table 1 correspond to the five and the one percent level of significance given that the correct r is used (i.e. $r = \rho$) . For $r = 0$, the values in the Table 1 represent $\Pr(\tilde{\underline{F}} \leq q)$, where $\tilde{\underline{F}}$ is given in (3.4). When $r \neq \rho$, the values in the table are those of the distribution function of $\hat{\underline{F}}$ in (3.7) . Notice also that the transformation using $r \neq \rho$ leads to a regression model with ARMA (1 , 1) errors.
From the results reported in Table 1 (and some additional unpublished results), it is apparent that the probability distribution can move considerably to the right, when the autocorrelation in the disturbances is underestimated $(r < \rho)$.

Table 1. The exact probability of not rejecting H_0, given that H_0 is true.

q	ρ = 0.0	ρ = 0.3	ρ = 0.6	ρ = 0.9	r	Restriction Matrix R	Time (s)
4.6	.950	.873	.762	.681	0.0	$\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}$	1
8.86	.990	.959	.890	.828	0.0		
4.6	.987	.950	.869	.792	0.3		1.199
8.86	.998	.990	.958	.914	0.3		
4.6	.995	.983	.950	.909	0.6		1.132
8.86	1.000	.998	.990	.976	0.6		
4.6	.993	.985	.968	.950	0.9		1.143
8.86	1.000	.999	.995	.990	0.9		
3.74	.950	.831	.605	.339	0.0	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	1.433
6.51	.990	.942	.784	.499	0.0		
3.74	.991	.950	.795	.488	0.3		1.323
6.51	.999	.990	.923	.672	0.3		
3.74	.998	.990	.950	.776	0.6		1.183
6.51	1.000	.999	.990	.914	0.6		
3.74	.993	.988	.975	.950	0.9		1.180
6.51	1.000	.999	.997	.990	0.9		
3.344	.950	.793	.466	.108	0.0	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	2.013
5.56	.990	.926	.670	.192	0.0		
3.344	.994	.950	.713	.195	0.3		
5.56	.999	.990	.878	.325	0.3		
3.344	1.000	.995	.950	.474	0.6		
5.56	1.000	1.000	.990	.661	0.6		
3.344	.998	.996	.990	.950	0.9		
5.56	1.000	1.000	.999	.990	0.9		
4.6	.950	.858	.676	.409	0.0	$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$	
8.86	.990	.951	.822	.543	0.0	$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$	
4.6	.950	.869	.749	.656	0.0	$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$	
8.86	.990	.957	.880	.807	0.0	$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$	
3.74	.950	.831	.605	.338	0.0	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$	
6.51	.990	.942	.783	.497	0.0		

The distribution moves to the left when the autocorrelation is overestimated $(r > \rho)$. The true probability of an error of type I increases substantially when $\rho = .9$ is severely underestimated. Also, the results seem to become worse as the number of restrictions increases.

In Table 2 , we shall compare the exact probability of an error of type I with the upper bounds (UB) on type I errors given by Kiviet (1979).

Table 2. Comparison of $Pr\,(UB_i \leq \tilde{F} \mid H_0)$ and $Pr\,(F_{i,14,\alpha=.05} \leq \tilde{F} \mid H_0)$

$i = 1\,,\,2$

number of restrictions: i	ρ	r	$P_r\,(UB_i \leq \tilde{F} \mid H_0)^*$	$Pr\,(4\,,6 \leq \tilde{F} \mid H_0)$
1	.3	.0	.16	.13 , .14 , .13
2	.3	.0	.19	.17 , .17
1	.9	.0	.63	.32 , .59 , .34
2	.9	.0	.73	.66 , .66
1	.0	.3	.12	.01
1	.0	.6	.17	.005
1	.0	.9	.19	.01
1	.3	.6	.10	.02
1	.3	.9	.12	.015
1	.6	.3	.21	.13
1	.6	.9	.08	.03
1	.9	.3	.52	.21
1	.9	.6	.31	.09

*) The upperbounds (UB) have been calculated by Kiviet (1979) for a sample size $T = 15$. We use a sample size $T = 17$.

The figures in column 5 of Table 2 correspond to those given in Table 1. From Table 2, it may be concluded that the exact probabilities, which depend on the specific regressor matrix and the restrictions, sometimes differ substantially from the upperbound probabilities.
Obviously, Kiviet's results and the figures presented in Table 1 indicate that care must be taken if the usual F- and t-tests are used in models with autocorrelated disturbances and when it is probable that the autocorrelation coefficient is estimated rather inaccurately. These findings clearly indicate that - contrary to widely accepted beliefs - the presence of autocorrelation in the disturbances inaccurately taken into account has much more serious consequences for the results of an empirical analysis than only a loss of efficiency.

Additional results not reproduced here show that the effects of first order moving average disturbances or the presence of heteroscedasticity on the probability of an error of type one are less important. In these situations, t- and F-tests may be more robust with respect to the properties of the disturbances. Of course, the results obtained for one data set need not to hold in general.

In the next section, we shall investigate the exact tail area probabilities for sample autocorrelation coefficients in ARIMA models and compare the exact probabilities with those associated with some large sample tests used in empirical work.

4. The exact distribution of the sample autocorrelations in ARIMA models

4.1. The distribution of the sample autocorrelation coefficients

Ratios of quadratic forms in normal variates frequently arise in econometrics, especially in the analysis of dynamic models. In the preceding section, the dynamics were introduced in the model through the presence of autocorrelated disturbances. In this section, we shall analyze the exact distribution of the k-th sample autocorrelation coefficient for a series of observations that is normally distributed and generated by an ARMA or an ARIMA model.

In order to briefly indicate how the exact distribution of sample autocorrelations can be obtained, we define the k-th sample autocorrelation, $r_k^{(T)}$, for a time series realization of size T, $\{y_1, y_2, \ldots, y_T\}$, as

$$r_k^{(T)} = \frac{\sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^{T} (y_t - \bar{y})^2} , \qquad (4.1)$$

where $\bar{y} = \frac{1}{T} \sum_{t=1}^{T} y_t$ is the sample mean. Using the lag operator L, an ARIMA (p, d, q) model for $\underline{y}_t$ can be written as

$$(1 - \phi_1 L - \ldots \phi_p L^p)(1 - L)^d \, \underline{y}_t = (1 - \theta_1 L - \ldots \theta_q L^q) \, \underline{a}_t , \qquad (4.2)$$

where the $\underline{a}_t$'s are assumed to be independent and normally distributed with mean zero and variance σ_a^2. We also assume that the polynomial $(1 - \phi_1 L - \ldots \phi_p L^p)$ has its roots outside the unit circle and the polynomial $(1 - \theta_1 L - \ldots \theta_q L^q)$ has no roots inside the unit circle (see e.g. Box and Jenkins (1970)). For the sake of simplicity, we assume that the mean of $\underline{y}_t$ is zero and $\sigma_a^2 = 1$.

Define the matrix $M = I_T - \frac{1}{T} \iota\iota'$, where ι is a $T \times 1$ vector of ones, and the matrix K_k, with all elements zero except those on the k-th upper- and lower-diagonal, which are equal to one. The sample autocorrelation coefficient in (4.1) can be written as:

$$\underline{r}_k^{(T)} = \frac{\underline{y}'\ M'\ K_k\ M\ \underline{y}}{2\ \underline{y}'\ M'\ M\ \underline{y}} \quad , \tag{4.3}$$

where $\underline{y} = (\underline{y}_1, \ldots, \underline{y}_T)'$. As the sample autocorrelation coefficient (4.3) is a RQFNV , Imhof's procedure presented in section 2 can be used to compute the exact distribution of $\underline{r}_k^{(T)}$, provided the covariance matrix of $\underline{y}$ is given.
First, we consider the case where $\underline{y}$ is generated by an ARMA (p , q) model (d = 0) . The covariance matrix of $\underline{y}$ exists and is constant.
In order to get the covariance matrix of $\underline{y}$ from the parameters of the ARMA model, it is convenient (see e.g. De Gooijer (1980)) to write the process as an infinite MA

$$\underline{y}_t = \underline{a}_t + \psi_1\ \underline{a}_{t-1} + \psi_2\ \underline{a}_{t-2} + \cdots \tag{4.4}$$

and for a sample of T observations as:

$$\underline{y} = \psi_\infty\ \underline{a}_\infty \quad , \tag{4.5}$$

where ψ_∞ is an $(T \times \infty)$ matrix and $\underline{a}_\infty$ is an $(\infty \times 1)$ column vector.
From (4.5) , the covariance matrix of $\underline{y} \sim N(0, \Sigma)$, becomes

$$\Sigma = \psi_\infty\ \psi_\infty' \quad , \tag{4.6}$$

so that the k-th sample autocorrelation coefficient in (4.3) can be written as a RQFNV

$$\underline{r}_k^{(T)} = \frac{\underline{\varepsilon}'\ C'\ M'\ K_k\ M\ C\ \underline{\varepsilon}}{2\ \underline{\varepsilon}'\ C'\ M'\ M\ C\ \underline{\varepsilon}} \quad , \tag{4.7}$$

where $\underline{\varepsilon} = C^{-1}\underline{y} \sim N(0, I)$, with C being a lower triangular matrix such that $\Sigma = C\ C'$.

When $p = 0$ in (4.2) , expression (4.5) specializes to

$$\underline{y} = \psi_{T+q}\ \underline{a}_{T+q} \quad , \tag{4.8}$$

where ψ_{n+q} is a $T \times (T+q)$ matrix and $\underline{a}_{T+q}$ is a $(T+q) \times 1$ column vector. When also $q = 0$, i.e. the process $\underline{y}$ is white noise, the matrices in the numerator and the denominator of (4.7) commute with each other, so that the characteristic roots corresponding to (4.7) are easily found for every value of the argument of the exact distribution of $\underline{r}_k^{(T)}$, once they have been computed for one non-zero value of the argument.

As $\underline{y}_t$ generated by an ARIMA (p , 1 , d) model is non-stationary, we express it in deviation from an initial value y_0 ,

$$\underline{z}_t = \underline{y}_t - y_0 = \sum_{i=0}^{t-1} [L^i \psi(L) \underline{a}_t] \quad , \tag{4.9}$$

where $\psi(L) = 1 + \psi_1 L + \psi_2 L^2 + \ldots$ is the polynomial of the infinite MA-representation of $(\underline{y}_t - \underline{y}_{t-1})$.

The process $\underline{z}_t$ can be represented as:

$$\underline{z} = \underline{y} - y_0 \iota = F \psi_\infty \underline{a}_\infty \quad , \tag{4.10}$$

where F is a $T \times T$ upper triangular matrix with all elements on and above the main diagonal equal to one. Conditionally on y_0 , the vector $\underline{y}$ is distributed as $N(y_0 \iota, \Sigma)$, with

$$\Sigma = F \psi_\infty \psi'_\infty F' \; . \tag{4.11}$$

The vector $\underline{z}$ is distributed as $N(0, \Sigma)$. As $\underline{z} - \bar{\underline{z}}\iota = \underline{y} - \bar{\underline{y}}\iota$, the k-th sample correlation coefficient for y_t as defined in (4.1) can be written as is done in (4.7) , with C obtained from Σ in (4.11). Clearly, it is possible to unify the computation of the exact distribution of the sample autocorrelations for ARMA (p, q) and ARIMA (p, 1, q) models. Obviously, the approach adopted for the ARIMA (p , 1 , q) model can also be extended for the cases where the integer $d > 1$. We shall now present some numerical results for the exact distribution of some sample autocorrelation coefficients, when the data are generated by an ARMA or an ARIMA model.

4.2. Some numerical results for the sample autocorrelations

In this subsection, we report results on the distribution of the first, second and fifth sample autocorrelation coefficients for the following models: two first order autoregressive (AR) models with $\phi = .9$ and $-.9$ respectively, two first order moving average (MA) models with $\theta = .7$ and $-.7$ respectively, two ARMA (1 , 1) models with $\phi = .9$, $\theta = .6$ and $\phi = .9$, $\theta = .3$ respectively, one integrated-moving average [IMA (1 , 1)] model, with $\theta = .6$ and one random walk model. The sample size is fixed at $T = 30$. We have chosen the first and second autocorrelations because they are usually important for the identification of a time series model. Results on $\underline{r}_5$ have been computed in order to give an indication of the shape of the distribution of a higher order autocorrelation.

For each sample autocorrelation coefficient, we also report a large sample normal approximation

$$\underline{r}_k \sim N\,[E\,(\underline{r}_k)\,,\,\mathrm{var}\,(\underline{r}_k)] \quad , \quad k = 1,\ 2,\ 5 \quad , \tag{4.12}$$

where the mean and the variance have been computed using an approximation correct to order T^{-1} for the expectation and the variance of $\underline{r}_k$ [see Kendall (1954)] .

De Gooijer (1980) gives the following approximate expressions for the mean and the variance of the sample autocorrelations:

$$E\,[\underline{r}_k^{(T)}] \simeq \frac{\mathrm{tr}\,(\Omega\, M_{k,d})}{\mathrm{tr}\,(\Omega\, N_d)}\left[1 - \frac{2\mathrm{tr}\,(\Omega\, M_{k,d}\,\Omega\, N_d)}{\mathrm{tr}\,(\Omega\, M_{k,d})\,\mathrm{tr}\,(\Omega\, N_d)} + \frac{2\mathrm{tr}\,(\Omega\, N_d)^2}{\mathrm{tr}^2\,(\Omega\, N_d)}\right] \tag{4.13}$$

where $\Omega = \psi\,\psi'$, $M_{k,d} = \frac{1}{2}\, F_d'\, M'\, K_k\, M\, F_d$, $N_d = F_d'\, M\, F_d$

with $F_d = I$ for $d = 0$ and $F_d = F$, $d = 1$,

and

$$\begin{aligned}\mathrm{var}\,(\underline{r}_k^{(T)}) \simeq \left(\frac{\mathrm{tr}\,(\Omega\, M_{k,d})}{\mathrm{tr}\,(\Omega\, N_d)}\right)^2 \Bigg[&\frac{2\mathrm{tr}\,(\Omega\, M_{k,d})^2}{\mathrm{tr}^2\,(\Omega\, M_{k,d})} + \frac{2\mathrm{tr}\,(\Omega\, N_d)^2}{\mathrm{tr}^2\,(\Omega\, N_d)} \\ &- \frac{4\mathrm{tr}\,(\Omega\, M_{k,d}\,\Omega\, N_d)}{\mathrm{tr}\,(\Omega\, M_{k,d})\,\mathrm{tr}\,(\Omega\, N_d)} - \left(\frac{2\mathrm{tr}\,(\Omega\, M_{k,d}\,\Omega\, N_d)}{\mathrm{tr}\,(\Omega\, M_{k,d})\,\mathrm{tr}\,(\Omega\, N_d)} \\ &- \frac{2\mathrm{tr}\,(\Omega\, N_d)^2}{\mathrm{tr}^2\,(\Omega\, N_d)}\right)^2\Bigg] .\end{aligned} \tag{4.14}$$

Expressions (4.13) and (4.14) have been used in the large sample normal approximation to the distribution of the sample autocorrelations reported in the Tables 3 to 10. At the bottom of each table, we report the exact value of ρ_k , the approximate value of the mean and variances as obtained from formulae (4.13) and (4.14), the value of the large sample variance according to Bartlett's (1946) formula for the asymptotic variance and the execution time (in seconds). For the non-stationary integrated models, no ρ_k and no value for Bartlett's variance are reported.
Notice also that the order in the sequence and the number of x , for which the distribution function has been evaluated, depends on the value of the distribution itself.

Table 3 . Distribution function of $\underline{r}_k$

Model : $\underline{y}_t = .9\ \underline{y}_{t-1} + \underline{a}_t$

x	Pr ($\underline{r}_1 \leq x$)	Norm. appr.	x	Pr ($\underline{r}_2 \leq x$)	Norm. appr.	x	Pr ($\underline{r}_5 \leq x$)	Norm. appr.
0.0000000	-.00006	.00000	0.0000000	.01586	.00142	0.0000000	.33192	.39713
.2500000	.00501	.00000	.2500000	.11644	.07415	-.2500000	.07056	.09750
.5000000	.06998	.01376	.0625000	.02813	.00467	-.0500000	.26434	.31996
.3333333	.01285	.00004	.1250000	.04733	.01338	-.1000000	.20338	.24989
.4166667	.03091	.00108	.1875000	.07589	.03361	-.1500000	.15010	.18895
.7500000	.51647	.65107	.5000000	.44159	.53709	-.2000000	.10564	.13810
.5416667	.10271	.03820	.3000000	.15933	.12759	-.5000000	.00298	.00987
.5833333	.14804	.09013	.3500000	.21279	.20324	-.3333333	.03145	.05040
.6250000	.20934	.18196	.4000000	.27763	.30071	-.4166667	.01129	.02352
.6666667	.28995	.31708	.4500000	.35410	.41503	.2500000	.72192	.78065
.7083333	.39231	.48252	.7500000	.92093	.94864	.0500000	.40519	.47857
1.0000000	.99999	.99856	.5416667	.52169	.63666	.1000000	.48307	.56091
.7916667	.65744	.79397	.5833333	.60638	.72774	.1500000	.56373	.64070
.8333333	.80155	.89477	.6250000	.69297	.80578	.2000000	.64490	.71476
.8750000	.92299	.95394	.6666667	.77772	.86841	.5000000	.97891	.96482
.9166667	.98927	.98284	.7083333	.85572	.91549	.3000000	.79089	.83680
.9583333	.99999	.99459	1.0000000	1.00001	.99924	.3500000	.85142	.88267
			.8333333	.99260	.98401	.4000000	.90402	.91856
			.9166667	1.00001	.99606	.4500000	.94737	.94548

k	1	2	5
ρ_k	.9	.81	.590
E ($\underline{r}_k$)	.71256	.48487	.06298
Var ($\underline{r}_k$)	.00930	.2640	.05833
Bartlett	.006	.022	.091
Time	2.380	2.742	2.9

Table 4. Distribution function of $\underline{r}_k$

Model : $\underline{y}_t = -.9\ \underline{y}_{t-1} + \underline{a}_t$

x	Pr ($\underline{r}_1 \leq x$)	Norm. appr.	x	Pr ($\underline{r}_2 \leq x$)	Norm. appr.	x	Pr ($\underline{r}_5 \leq x$)	Norm. appr.
0.0000000	1.00026	1.00000	0.0000000	.00263	.00000	0.0000000	.90918	.93028
-.2500000	.99961	1.00000	.2500000	.02639	.00025	-.2500000	.69792	.65295
-.5000000	.98735	1.00000	.0833333	.00609	.00000	-.0625000	.87069	.88623
-.3333333	.99841	1.00000	.1666667	.01304	.00001	-.1250000	.82264	.82526
-.4166667	.99532	1.00000	.5000000	.15968	.08734	-.1875000	.76470	.74680
-.7500000	.79981	.86656	.3125000	.04323	.00158	-.5000000	.36060	.24470
-.5625000	.97390	.99998	.3750000	.06875	.00775	-.3000000	.63896	.57001
-.6250000	.94710	.99904	.4375000	.10622	.02945	-.3500000	.57524	.48384
-.6875000	.89548	.98240	.7500000	.61333	.77917	-.4000000	.50698	.39842
-1.0000000	.00001	.00203	.5416667	.20644	.15797	-.4500000	.43487	.31763
-.7812500	.72688	.72985	.5833333	.26376	.25836	-.7500000	.05280	.03788
-.8125000	.63197	.54557	.6250000	.33283	.38439	-.5500000	.28728	.18190
-.8437500	.51225	.35070	.6666667	.41445	.52412	-.6000000	.21895	.13028
-.8750000	.36878	.18907	.7083333	.50852	.66090	-.6500000	.15680	.08980
-.9062500	.21244	.08392	1.0000000	1.00001	.99811	-.7000000	.10060	.05951
-.9375000	.07364	.03025	.8000000	.74671	.88390	-1.0000000	.00002	.00212
-.9687500	.00404	.00877	.8500000	.87300	.94739	-.8333333	.00495	.01629
			.9000000	.96570	.97959	-.9166667	.00002	.00623
			.9500000	.99939	.99326	.2500000	.98738	.99480
						.0833333	.94722	.96707
						.1666667	.97242	.98613

k	1	2	5
ρ_k	-.9	.81	-.590
E ($\underline{r}_k$)	-.81968	.65956	-.34066
Var ($\underline{r}_k$)	.00394	.01382	.05313
Bartlett	.006	.022	.091
Time	2.534	2.847	3.086

Table 5. Distribution function of $\underline{r}_k$

Model : $\underline{y}_t = \underline{a}_t - .7\ \underline{a}_{t-1}$

x	Pr ($\underline{r}_1 \leq x$)	Norm. appr.	x	Pr ($\underline{r}_2 \leq x$)	Norm. appr.	x	Pr ($\underline{r}_5 \leq x$)	Norm. appr.
0.0000000	.99867	.99956	0.0000000	.53106	.53015	0.0000000	.50418	.50405
-.2500000	.91855	.92584	-.2500000	.12276	.13345	-.2500000	.09509	.10679
-.0833333	.99536	.99665	-.0500000	.43486	.43584	-.0500000	.40151	.40491
-.1666667	.97391	.98116	-.1000000	.34174	.34506	-.1000000	.30490	.31156
-.5000000	.34529	.32534	-.1500000	.25645	.26243	-.1500000	.21966	.22898
-.2916667	.86768	.87057	-.2000000	.18271	.19131	-.2000000	.14929	.16035
-.3333333	.79667	.79181	-.5000000	.00462	.01083	-.5000000	.00265	.00625
-.3750000	.70439	.69017	-.3125000	.06790	.07975	-.3125000	.04881	.05970
-.4166667	.59328	.57140	-.3750000	.03315	.04427	-.3750000	.02192	.03069
-.4583333	.47007	.44574	-.4375000	.01390	.02277	-.4375000	.00840	.01447
-.7500000	.00272	.00936	.2500000	.90322	.89645	.2500000	.90866	.89690
-.5500000	.21048	.20257	.0500000	.62523	.62280	.0500000	.60659	.60293
-.6000000	.10798	.11273	.1000000	.71260	.70885	.1000000	.70248	.69558
-.6500000	.04428	.05572	.1500000	.78939	.78442	.1500000	.78665	.77712
-.7000000	.01348	.02434	.2000000	.85320	.84717	.2000000	.85575	.84454
			.5000000	.99613	.99281	.5000000	.99757	.99410
			.3125000	.94739	.94039	.3125000	.95351	.94267
			.3750000	.97449	.96817	.3750000	.97932	.97069
			.4375000	.98921	.98426	.4375000	.99217	.98626

k	1	2	5
ρ_k	-.470	0.0	0.0
E ($\underline{r}_k$)	-.44037	-.01595	-.00202
Var ($\underline{r}_k$)	.01734	.04444	.03975
Bartlett	.018	.048	.048
Time	1.868	2.336	2.350

Table 6. Distribution function of $\underline{r}_k$

Model : $\underline{y}_t = \underline{a}_t + .7\, \underline{a}_{t-1}$

x	Pr ($\underline{r}_1 \leq$ x)	Norm. appr.	x	Pr ($\underline{r}_2 \leq$ x)	Norm. appr.	x	Pr ($\underline{r}_5 \leq$ x)	Norm. appr.
0.0000000	.00338	.00158	0.0000000	.65243	.64581	0.0000000	.61802	.61283
.2500000	.13613	.13429	-.2500000	.19301	.20270	-.2500000	.14607	.15550
.0625000	.01038	.00638	-.0416667	.57419	.56869	-.0416667	.53057	.52792
.1250000	.02782	.02123	-.0833333	.49251	.48884	-.0833333	.44156	.44173
.1875000	.06853	.05848	-.1250000	.41046	.40944	-.1250000	.35508	.35822
.5000000	.75967	.77007	-.1666667	.33129	.33360	-.1666667	.27498	.28100
.2857143	.19557	.19969	-.2083333	.25798	.26402	-.2083333	.20431	.21287
.3214286	.26997	.28127	-.5000000	.01028	.02077	-.5000000	.00513	.01036
.3571429	.35823	.37621	-.3125000	.11462	.12850	-.3125000	.07888	.09043
.3928571	.45717	.47933	-.3750000	.06027	.07564	-.3750000	.03758	.04815
.4285714	.56157	.58387	-.4375000	.02731	.04123	-.4375000	.01528	.02340
.4642857	.66474	.68276	.2500000	.94919	.94296	.2500000	.95372	.94569
.7500000	.99910	.99912	.0500000	.73801	.73080	.0500000	.71549	.70769
.5625000	.88920	.88501	.1000000	.81144	.80415	.1000000	.79981	.79006
.6250000	.96280	.95172	.1500000	.87097	.86383	.1500000	.86778	.85692
.6875000	.99198	.98313	.2000000	.91649	.90970	.2000000	.91862	.90767
			.5000000	.99881	.99733			
			.3333333	.98107	.97627			
			.4166667	.99446	.99144			

k	1	2	5
ρ_k	.470	0.0	0.0
E ($\underline{r}_k$)	.39988	-.07753	-.05514
Var ($\underline{r}_k$)	.01835	.04297	.03699
Bartlett	.018	.048	.048
Time	1.995	2.305	1.989

Table 7. Distribution function of $\underline{r}_k$

Model : $\underline{y}_t = .9\ \underline{y}_{t-1} + \underline{a}_t - .6\ \underline{a}_{t-1}$

x	Pr ($\underline{r}_1 \leq x$)	Norm. appr.	x	Pr ($\underline{r}_2 \leq x$)	Norm. appr.	x	Pr ($\underline{r}_5 \leq x$)	Norm. appr.
0.0000000	.12130	.14232	0.0000000	.19058	.22227	0.0000000	.41909	.44550
-.2500000	.00974	.01572	-.2500000	.01711	.03105	-.2500000	.05007	.07500
-.0625000	.07356	.09007	-.0625000	.12001	.14922	-.0500000	.31386	.34548
-.1250000	.04121	.05363	-.1250000	.06939	.09425	-.1000000	.22127	.25526
-.1875000	.02109	.02999	-.1875000	.03642	.05588	-.1500000	.14581	.17917
.2500000	.47298	.50455	.2500000	.61451	.63174	-.2000000	.08911	.11919
.0500000	.17192	.19664	.0416667	.24857	.28061	-.5000000	.00052	.00305
.1000000	.23377	.26193	.0833333	.31424	.34548	-.3333333	.01546	.03049
.1500000	.30598	.33683	.1250000	.38597	.41526	-.4166667	.00545	.01050
.2000000	.38665	.41884	.1666667	.46160	.48782	.2500000	.88951	.87808
.5000000	.86472	.86275	.2083333	.53864	.56079	.0416667	.51162	.53189
.3000000	.56148	.59005	.5000000	.93713	.92471	.0833333	.60370	.61681
.3500000	.64838	.67146	.3000000	.70057	.71112	.1250000	.69083	.69644
.4000000	.72993	.74545	.3500000	.77782	.78138	.1666667	.76912	.76769
.4500000	.80287	.80962	.4000000	.84371	.84063	.2083333	.83579	.82852
.7500000	.99672	.98515	.4500000	.89687	.88825	.5000000	.99918	.99320
.5625000	.92435	.91358	.7500000	.99964	.99443	.3125000	.94638	.93202
.6250000	.96420	.94881	.5833333	.97839	.96442	.3750000	.97908	.96536
.6875000	.98677	.97153	.6666667	.99544	.98505	.4375000	.99429	.98391

k	1	2	5
ρ_k	.493	.444	.323
E ($\underline{r}_k$)	.24736	.17360	.02631
Var ($\underline{r}_k$)	.05345	.05156	.03684
Bartlett	.067	.075	.094
Time	2.625	2.633	2.604

Table 8. Distribution function of $\underline{r}_k$

Model : $\underline{y}_t = .9\ \underline{y}_{t-1} + \underline{a}_t - .3\ \underline{a}_{t-1}$

x	Pr ($\underline{r}_1 \leq x$)	Norm. appr.	x	Pr ($\underline{r}_2 \leq x$)	Norm. appr.	x	Pr ($\underline{r}_5 \leq x$)	Norm. appr.
0.0000000	.00683	.00047	0.0000000	.04400	.03307	0.0000000	.34896	.41020
.2500000	.06954	.03754	.2500000	.24553	.27329	-.2500000	.05518	.09495
.0833333	.01632	.00258	.0625000	.07376	.06316	-.0500000	.26866	.32860
.1666667	.03520	.01105	.1250000	.11620	.11120	-.1000000	.19749	.25448
.5000000	.33919	.39962	.1875000	.17317	.18101	-.1500000	.13756	.19016
.3000000	.10067	.07012	.5000000	.65562	.73619	-.2000000	.09014	.13696
.3500000	.14180	.12105	.3000000	.31420	.36092	-.5000000	.00117	.00832
.4000000	.19455	.19363	.3500000	.39141	.45657	-.3333333	.02044	.04725
.4500000	.26016	.28792	.4000000	.47545	.55481	-.4166667	.00576	.02101
.7500000	.85334	.89820	.4500000	.56390	.64978	.2500000	.77877	.80422
.5416667	.41499	.49998	.7500000	.97844	.96899	.0416667	.42097	.48150
.5833333	.49869	.60035	.5500000	.74084	.81017	.0833333	.49578	.55340
.6250000	.58824	.69446	.6000000	.82125	.86980	.1250000	.57138	.62358
.6666667	.68044	.77721	.6500000	.89027	.91502	.1666667	.64645	.68988
.7083333	.77075	.84544	.7000000	.94361	.94730	.2083333	.71539	.75053
1.0000000	1.00000	.99742				.5000000	.99133	.97385
.8125000	.94813	.95081				.3125000	.85884	.87029
.8750000	.99328	.97903				.3750000	.92126	.91905
.9375000	.99998	.99215				.4375000	.96585	.95251

k	1	2	5
ρ_k	.796	.717	.522
E ($\underline{r}_k$)	.54167	.37209	.05237
Var ($\underline{r}_k$)	.02685	.04101	.05321
Bartlett	.022	.040	.099
Time	2.622	2.118	2.644

Table 9. Distribution function of $\underline{r}_k$

Model : $\underline{y}_t = \underline{y}_{t-1} + \underline{a}_t - .6\ \underline{a}_{t-1}$

x	Pr ($\underline{r}_1 \le x$)	Norm. appr.	x	Pr ($\underline{r}_2 \le x$)	Norm. appr.	x	Pr ($\underline{r}_5 \le x$)	Norm. appr.
0.0000000	.06260	.07390	0.0000000	.10568	.13835	0.0000000	.28685	.35867
-.2500000	.00447	.00623	-.2500000	.00841	.01662	-.2500000	.03008	.06236
-.0833333	.03038	.03611	-.0625000	.06436	.08881	-.0500000	.20841	.27536
-.1666667	.01276	.01585	-.1250000	.03610	.05387	-.1000000	.14290	.20290
.2500000	.28582	.34601	-.1875000	.01842	.03082	-.1500000	.09177	.14322
.0625000	.09943	.11810	.2500000	.40603	.48154	-.2000000	.05476	.09669
.1250000	.14879	.17833	.0500000	.14914	.18958	.2500000	.74046	.79139
.1875000	.21115	.25497	.1000000	.20191	.25106	.0416667	.35996	.43389
.5000000	.65638	.74380	.1500000	.26329	.32175	.0833333	.43760	.51160
.3000000	.35316	.42627	.2000000	.33195	.39953	.1250000	.51706	.58886
.3500000	.42574	.50972	.5000000	.78250	.84018	.1666667	.59558	.66282
.4000000	.50178	.59275	.3000000	.48339	.56435	.2083333	.67068	.73096
.4500000	.57935	.67179	.3500000	.56180	.64442	.5000000	.99135	.97640
.7500000	.95751	.95603	.4000000	.63909	.71857	.3000000	.81549	.85220
.5500000	.73076	.80658	.4500000	.71327	.78433	.3500000	.87982	.89983
.6000000	.80037	.85895	.7500000	.99206	.97916	.4000000	.93191	.93514
.6500000	.86298	.90076	.5625000	.85937	.89536	.4500000	.96961	.95993
.7000000	.91621	.93269	.6250000	.92221	.93523			
			.6875000	.96730	.96216			

k	1	2	5
E ($\underline{r}_k$)	.34420	.2611	.07714
Var ($\underline{r}_k$)	.05655	.05762	.04540
Time	2.515	2.651	2.391

Table 10. Distribution function of r_k

Model : $y_t = y_{t-1} + a_t$

x	Pr ($r_1 \leq x$)	Norm. appr.	x	Pr ($r_2 \leq x$)	Norm. appr.	x	Pr ($r_5 \leq x$)	Norm. appr.
0.0000000	-.00024	0.00000	0.0000000	.00741	.00000	0.0000000	.20702	.17759
.2500000	.00221	0.00000	.2500000	.05919	.00064	-.2500000	.03895	.00837
.5000000	.03372	.00000	.0833333	.01589	.00000	-.0625000	.15020	.09825
.3333333	.00584	.00000	.1666667	.03175	.00003	-.1250000	.10299	.04860
.4166667	.01431	.00000	.5000000	.27067	.22698	-.1875000	.06588	.02140
.7500000	.32793	.35711	.3125000	.09065	.00461	.2500000	.53163	.70657
.5500000	.05528	.00006	.3750000	.13439	.02354	.0500000	.25956	.26402
.6000000	.08897	.00145	.4375000	.19337	.08579	.1000000	.31868	.36791
.6500000	.14034	.01751	.7500000	.79230	.95770	.1500000	.38441	.48253
.7000000	.21692	.10802	.5416667	.33374	.36822	.2000000	.45614	.59863
1.0000000	1.00000	.99997	.5833333	.40680	.53015	.5000000	.91977	.97786
.7857143	.43374	.60099	.6250000	.49012	.68719	.3000000	.60764	.79670
.8214286	.56438	.81003	.6666667	.58352	.81597	.3500000	.68424	.87089
.8571429	.71771	.93321	.7083333	.68565	.90530	.4000000	.76317	.92280
.8928571	.87637	.98309	1.0000000	1.00004	.99999	.4500000	.84351	.95708
.9285714	.98328	.99697	.8125000	.93592	.99043	.7500000	.99989	.99975
.9642857	1.00000	.99962	.8750000	.99821	.99847	.5833333	.99684	.99380
			.9375000	1.00003	.99985	.6666667	.99993	.99861

k	1	2	5
E (r_k)	.77102	.575569	.15746
Var (r_k)	.00330	.01022	.02900
Time	2.575	2.985	2.861

Most interesting is the comparison of the exact distribution with the normal approximation.
For autoregressive models, there may be substantial differences between the distribution functions, so that the normal approximation performs rather badly. For pure moving average models, the differences between the two distributions are much less pronounced. As one now expects for ARMA models, the large sample normal distribution is again less appropriate as an approximation in small sample situations.
A similar conclusion holds for the integrated models. These conclusions seem to depend to some extent on the sign of the parameters in the ARIMA-models.

It is also interesting to compare the expected value of $\underline{r}_k$ with the value of the population coefficient ρ_k .
As indicated by De Gooijer (1980) , the approximations for the expectation and the variance in (4.13) and (4.14) are very close to the exact expectation and variance of the $\underline{r}_k$'s . For pure autoregressive models, the value of Bartlett's formulae for the variance sometimes differs substantially from var $(\underline{r}_k)$. For the pure moving average models and for ARMA-models Bartlett's formulae takes values close to var $(\underline{r}_k)$.

Although we have analyzed the distribution of three sample autocorrelations for a few models only, we are tempted to conclude that the knowledge of the exact distribution may help in discriminating between alternative ARIMA-models. For the analysis of economic time series, which usually have a pronounced autoregressive shape, this conclusion may be very relevant.

It has also become clear that for the models analyzed in this section, the large sample normal distribution with mean and variance computed according to Kendall's (1954) formulae is sometimes a good approximation to the exact distribution of sample autocorrelations. Given that the computation of the approximate expressions for the mean and variance given in (4.13) and (4.14) is fairly straightforward, this approximation seems to be a reasonable and cheap alternative for the exact distribution. However, its properties deserve further investigation.

5. Some concluding remarks

In this paper, we have presented two applications of the exact distribution of ratios of quadratic forms in normal variates arising in some dynamic econometric models. Although Imhof's (1961) and other similar procedures for computing RQFNV's have been used on several occasions, their applicability does not seem to be fully appreciated. RQFNV's arise in many statistical and econometric problems. In the introduction, we referred to several applications in the literature. Imhof's procedure has and will be used to compute tables for statistical tests, e.g. the Durbin-Watson test, or to extend existing ones. The procedure will also be very useful for the comparison of the statistical techniques involving RQFNV's. In these circumstances, integration (Imhof) is a substitute for the simulation techniques where the exact distribution is obtained through sufficient sampling. The choice between integration or simulation will depend on the relative costs of implementation (see e.g. Kloek and Van Dijk (1978) for an application where sampling is preferred to integration).
We have used Imhof's procedure to obtain exact tail area probabilities for tests of linear restrictions on the regression coefficients of models in which the dynamics enter through the autocorrelation of the disturbances. Then we have evaluated the exact distribution of sample autocorrelation coefficients for some stationary and some non-stationary models and compared the exact distribution with some large sample approximations.

For the regression model, our results indicate that the exact distribution sometimes differs substantially from an incorrectly assumed F-distribution and from the lower and upper bounds for the critical region. Care has to be taken, even when the disturbance autocorrelation has been approximately corrected for.

For the ARMA models, but in particular for MA models, a large sample normal approximation based on Kendall's (1954) formulae for the mean and variance seems to be a reasonable alternative to the exact distribution of the sample autocorrelation coefficients.
Still in both applications it is recommendable to compute the exact distribution.
The implementation of Imhof's procedure has its limitations, especially when the number of normal variates (sample size) is large. For small samples, when the need for the exact distribution is largest, our results suggest that the computation of the exact distribution yields new and valuable insights for a reasonable price in terms of computational work.

References

Anderson, O.D. (1979): "Formulae for the Expected Values of the Sampled Variance and Covariances from Series Generated by General Autoregressive Integrated Moving Average Process of Order (p, d, q)", Sankhyā, forthcoming.

Bartlett, M.S. (1946): "On the Theoretical Specification of Sampling Properties of Autocorrelated Time Series", Journal of the Royal Statistical Society, B, 8, 27.

Box, G.E.P., and G.M. Jenkins (1970): Time Series Analysis, Forecasting and Control, San Francisco, Holden-Day.

De Gooijer, J.G. (1980): "Exact Moments of the Sample Autocorrelations from Series Generated by General ARIMA Process of Order (p, d, q), d = 0 or 1", Journal of Econometrics, 14, 365-379.

Dubbelman, C., Louter, A.S., and A.P.J. Abrahamse (1978): "On Typical Characteristics of Economic Time Series and the Relative Qualities of Five Autocorrelation Tests", Journal of Econometrics, 8, 295-306.

Durbin, J., and G.S. Watson (1950): "Testing for Serial Correlation in Least Squares Regression 1", Biometrika, 37, 409-428.

Farebrother, R.W. (1980): "The Durbin-Watson Test for Serial Correlation when there is no Intercept in the Regression", Econometrica, 48, 1553-1564.

Hammersley, J.M. and D.C. Handscomb (1964): Monte Carlo Methods, London, Methuen.

Imhof, J.P. (1961): "Computing the Distribution of Quadratic Forms in Normal Variates", Biometrika, 48, 419-426.

Kendall, M.G. (1954): "Note on Bias in the Estimation of Autocorrelation", Biometrika, 41, 403-404.

Kiviet, J.F. (1977): "Non-detection of serial correlation in least squares regression; frequency and consequences", Report, University of Amsterdam.

Kiviet, J. (1979): "Bounds for the Effects of ARMA Disturbances on Tests for Regression Coefficients", Report, Unversity of Amsterdam.

Kloek, T., and H.K. van Dijk (1978): "Bayesian Estimates of Equation Parameters: An Application of Integration by Monte Carlo", Econometrica, 46, 1-19.

Koerts, J. and A.P.J. Abrahamse (1969): On the Theory and Application of the General Linear Model, Rotterdam University Press, Rotterdam.

Rao, C.R. (1967): "Least Squares Theory Using an estimated Dispersion Matrix and Its Application to Measurement of Signals", Proceedings of the Fifth Berkeley Symposium, Vol. 1, Berkeley, University of California Press, 355-372.

Savin, N.E., and K.J. White (1977): "The Durbin-Watson Test for Serial Correlation with Extreme Sample Sizes and Many Regressors", Econometrica, 45, 1989-1996.

Theil, H. (1971): "Principles of Econometrics, New York, J. Wiley and Sons.